Application of emotion recognition methods in automotive research

Martin A. Tischler, Christian Peter, Matthias Wimmer, Jörg Voskamp

[1]Customer Research Center, DaimlerChrysler AG, Böblingen
[2]Fraunhofer Institut für Graphische Datenverarbeitung, Rostock
[3]Lehrstuhl für Informatik IX, Technische Universität München

Abstract. This paper reports on a pilot study applying emotion recognition technologies developed for Human-Machine-Interfaces in automobile research. The aim of the study was to evaluate technologies for quantifying driving pleasure in a close-to-reality scenario. Results show that car driving scenarios pose particular requirements on emotion recognition technologies which could be met by modifications of current systems.

Keywords: Driving Pleasure, Emotion Recognition, Affective Interaction, Human-Machine-Interaction, Affect Sensors, Facial Expression Interpretation

1 Introduction

A task in vehicle development is to configure vehicles not only according to security aspects, but also to ensure passengers feeling well and enjoying their drive. The pleasure of driving a car or, more general "the positive emotional interaction with the product car" is vitally important for the acceptance and the success of a vehicle model today [1]. Driving pleasure is determined particularly by the driving feeling, which arises from the complex reciprocal relationship driver-vehicle-environment.
Measuring emotional reactions of drivers in close-to-reality environments opens extended views of the interaction driver-vehicle and supplies important information to design vehicles to better comply to and satisfy constantly increasing customer demands.

This paper reports on a pilot study applying emotion recognition technologies developed for Human-Machine-Interfaces in automobile research. The aim of the study was to evaluate technologies to quantify driving pleasure in a close-to-reality scenario.

2 A Short Overview of Research Findings on Driving Pleasure

The topic “driving pleasure” was primarily explored by means of interviews or studies in driving simulators. Only very few studies have been performed to capture positive experiences of non-professional drivers in natural environments [2]. Beyond that no scientific investigation is known assessing driving fun of non-professional drivers with quantitative measures. While common studies rely solely on retrospective questionnaires and interviews, for product development purposes new methods to quantify the emotional state of the driver are needed. They would provide objective information in very small time intervals without interference of the driver. This accurate information is desirable, because in future research it is intended to combine objective conditions for instance driving dynamic measures with the emotional reaction of the driver. It is remarkable that the positive aspects of the activity "driving a car" were neglected so far by psychological research, although studies showed consistently that the majority of the people enjoy the activity driving [3].

3 Technologies Applied

Emotions are manifest in physiological changes controlled by the autonomous nervous system [4]. They can either be observed by humans, such as facial features, gestures, or changes of the voice, or they can be less obvious, like heart rate, blood pressure, skin temperature, electro-dermal activity, or pupil dilation. Sensors and modern computing technology today offer possibilities to access a variety of emotion-related parameters, including those not observable by humans, in a non- or minimally intrusive way. In the following we shortly describe the technologies used to infer the emotional state of drivers.

Emotion Recognition from Speech: Speech parameters have been examined for correlations with emotions for a long time. In connection with computers and other applications, results are increasingly acceptable [5].

At Fraunhofer IGD we work on tools to extract emotion information from a user's voice. While our research covers quality features of the speech signal, prosody, and semantics, in this study only emotion-related features of the speech signal have been analysed [6, 7]. For analysing the data, techniques from data mining and knowledge discovery have been applied. The extracted information then is used to build two classifiers based on the valence/arousal model [cf. 8, 9], one for detecting valence and one for arousal. They were trained on 10% of the collected data and tested using 10-fold cross-validation. They were then applied on the remaining 90%.

Emotion Recognition from Facial Features: During the last decade, a lot of research on facial expression recognition has been conducted [10]. Model-based approaches have proven to be highly appropriate for interpreting images in real-world scenarios. At the Technische Universität München, we developed algorithms that robustly localize facial features and seamlessly track them through image sequences in order to interpret facial expressions. They are able to determine the six universal emotions specified by Ekman [11], and to give evidence about the magnitude of the visible expression.

An Active Shape Model [12] represents the contour of the face and it is fit to the camera image via learned objective functions [13]. Successfully tracking a human face during some period of time, our system determines the motion of the individual parts such as eyes, chin, cheeks, and lips. A previously learned classifier uses this information to infer the currently visible facial expression.

Emotion Recognition from Physiological Data: Emotion-related changes of physiological parameters have been studied for a long time [14] and can be considered to be the emotional signs best understood today. A number of proof-of-possibility studies for emotion detection in non-lab settings have been performed, using either commercially available sensor equipment for physiology data or experimental devices. Among the latter are stationary, mobile, or wearable systems [15, 16], or furniture equipped with sensors [17].

At Fraunhofer IGD we developed a wireless and easy to use sensor system for collecting the emotion-related physiological parameters skin resistance, skin temperature, and heart rate [16]. As can be seen in figure 1, the EREC system (for *E*motion *REC*ognition) consists of a glove hosting sensing elements for skin temperature and skin resistance, a wrist pocket containing sensing electronics, and a base unit performing data enhancement steps and storing the data on an SD card. The system is particularly suitable for mobile applications since it provides for both, immediate data communication to a nearby computer, and storage of data on an exchangeable memory card. It has been used in several studies of different application fields, inside and outside a lab, and proved to be stable and reliable [18]. In this study, the wireless option has proved to be very useful for checking the quality of data in the setup phase. During the study, data have been recorded on memory card.

4 Study

The objective of this pilot study was to apply sophisticated emotion recognition techniques into a test car and to demonstrate their applicability for the described purpose. Three research institutes cooperated and provided two cars with sensory equipment to conduct a study in a close-to-reality environment. For safety reasons it was decided to do the study on a separated proving ground in Southern Germany.

A middle class saloon car (‘new car’) with a performance of 170 hp and a 25 year old previous model of this series (‘forerunner’) were used for a comparative study. Eight participants (age 33-53), all non-professional drivers, were invited. They drove on their own on three different courses: an autobahn-like course, a course similar to a rural road and a demanding handling course. Figure 1 shows the technical setup.

4.1 Setup

Technical setup: For recording the speech data, a capacitor lapel microphone has been connected to a Marantz PMD 660 audio recorder which stored the data on a removable flash card. The microphone was attached to the safety belt, at chest level. The recorder was placed on the backseat.

The EREC glove and chest belt were put on by the test persons at the beginning of the session. Sufficient time was allowed for the users to become familiar with the components and adjust them for most possible comfort. The base unit was placed in the center storage shelf of the car. For the facial data, a small video camera was fixed at the front column of the car. The video signal was recorded by a Sony Digital Video Walkman, which was placed on the backseat.

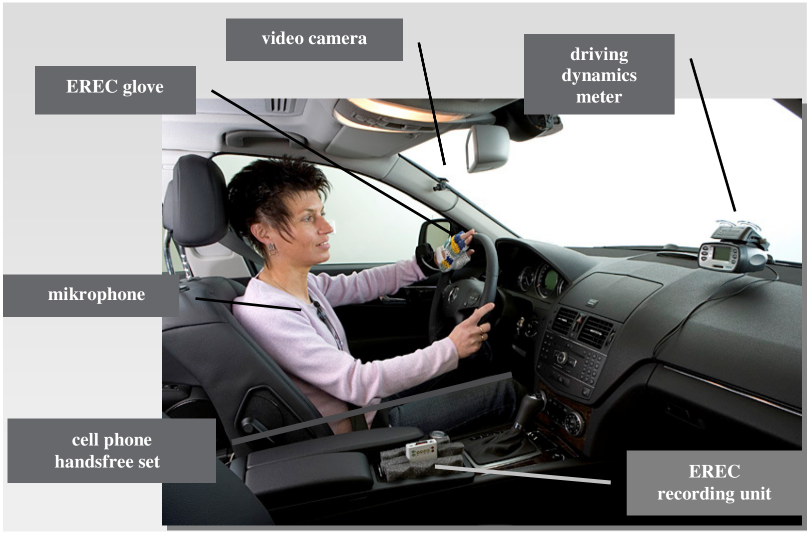

Fig 1. Technical setup of the study

Both cars were equipped with cell phones and hands free sets, in order to give instructions to the drivers. For logging the actual position and the basic driving dynamic data the GPS-based Racelogic VBOX Mini was used.

Additional data acquisition: Additional to the quantitative measurements of the emotional state, the participants were asked via questionnaires and in interviews before, during, and after driving about their subjective feelings. These results and the detailed methods are not presented here.

4.2 Results

Speech analysis: Parameters found to be most related with emotional speech were changes of pitch, intensity, and energy changes over frequency bands. We achieved an average confidence of 1.1-1.2 for valence and 0.9-1.0 for arousal.

Valence is found to be generally higher for the new car (6.11 vs. 5.45), particularly so for the handling course (6.62 vs. 5.58). Arousal is just slightly higher for the new car (6.0. vs. 5.95) but more significantly at the handling course again (6.13 vs. 5.47).

Facial features: Our system recognizes well-posted facial expressions that incorporate a lot of facial motion best, such as happiness or surprise. In these cases, facial motion is well-distinguished from head motion and background motion. In the car-scenario, we are interested in determining happiness, however, the results of our investigation show that drivers don't express this emotion strongly while being alone in the car. The presence of non emotional-related head motion while observing the traffic makes facial expression recognition very difficult.

The results of our investigation illustrate that the drivers express happiness with a magnitude of 5.5% on average and with a peak value of 19.8% in the new car. In contrast, happiness is expressed with a magnitude of 4.2% in the forerunner car (15.0% peak value). The underlying scale is derived from the image sequences of the Cohn-Kanade-Facial-Expression-Database [19], where the first image denotes a neutral face (0%) and the last image denotes the maximum expression (100%).

Physiology: Physiological data have been gathered with different reliability. While skin resistance and skin temperature proved to be reliable data sources, heart rate data from the chest belt were less secure. As elaborated in consecutive tests, the chest belt data link was most likely affected by electro-magnetic interferences from servo and wiper drives. Further we found that the amount of training data for classifiers was too small to achieve results of sufficient confidence. Due to these circumstances we decided to not take physiological data into account for the final results.

5 Discussion

Speech data reveals that in the forerunner car arousal rises significantly on the speedway, compared to the other courses. In combination with results from interviews this leads to the conclusion that drivers of the forerunner car feel more fearful at this part of the course. With the new car arousal is higher on the speedway as well. Interview results lead to the conclusion that this increase in arousal is due to feelings of happiness, fun and joy.

Recognizing facial expressions in real world scenarios poses a higher challenge to both correctly fitting the face model and classifying the facial expression. In lab environments colour, texture, and motion of the background as well as the brightness and the position of the light source are controllable. While showing facial expressions in the laboratory, people are often asked to stop talking and reduce head movements as well as facial muscle activity that are not related to facial expressions. However, these activities occur often in real world scenarios and therefore, they have to be taken into account.

For the physiological data it can be said that other means for detecting heart rate needs to be found for future studies in cars. The main reason for the physiological data not being usable in this study was the lack of sufficient training data. This was mainly due to the tight schedule which didn't allow for an extra data collection session with each participant. Our plans to retrieve sufficient training data at the beginning of the main session proved to be unsuitable.

Concluding it can be said that all modalities give evidence that valence, i.e. the feeling well of the test persons was generally much higher in the new car. This study can also be seen as a prove that affect sensors can be applied to measure positive emotions in real-world settings, and that only a combined analysis of all modalities' results leads to useful results.

References

1. Automobilwoche, 23.4. 2007, p. 12
2. Toda, Y., & Kageyama, I.: Measuring a Driver's Pleasure during Driving. Rev. of Automotive Engineering 28, 89--93 (2007)
3. Ellinghaus, D., Steinbrecher, J.: Fahren auf Landstraßen. Traum oder Albtraum? Untersuchungen zum Fahrverhalten und Fahrvergnügen von Pkw-Fahrern auf Landstraßen. Uniroyal Verkehrsuntersuchung 28. Köln/Hannover (2003)
4. Lane, R.D., Nadel L. (eds.): Cognitive Neuroscience of Emotion. Oxford University Press (2000)
5. Douglas-Cowie, E., Cowie, R., Campbell, N.: Speech and Emotion. Speech Communication 40, 1--3 (2003)
6. Fernandez, R.: A Computational Model for the Automatic Recognition of Affect in Speech. Ph.D. thesis, Massachusetts Institute of Technology (2004)
7. Lee, C.M., Narayanan, S.S.: Toward detecting emotions in spoken dialogs. IEEE Transactions on Speech and Audio Processing, 13(2) (2005)
8. Picard R.W., Vyzas E., Healey J.: Toward Machine Emotional Intelligence - Analysis of Affective Physiological State. IEEE Trans. on Pattern Analysis and Machine Intelligence, 23(10) (2001)
9. Russell, J. A.: A Circumplex Model of Affect. Journal of Personality and Social Psychology 39, 161--1178 (1980)
10. Cohen, I., Sebe, N., Chen, L., Garg, A., Huang, T.: Facial expression recognition from video sequences: Temporal and static modeling. CVIU, 91(1--2), 160—187 (2003)
11. Ekman, P.: Universals and Cultural Differences in Facial Expressions of Emotion. In: Cole, J. K. (ed.), Nebraska Symposium on Motivation, vol. 19. Lincoln: University of Nebraska Press (1971)
12. Cootes, T.F., Taylor, C.J.: Active shape models - smart snakes. In: Proc. of the 3rd British Machine Vision Conference 1992, pp. 266--275. Springer (1992)
13. Wimmer, M., Pietzsch, S., Stulp, F., Radig, B.: Learning robust objective functions with application to face model fitting. 29th DAGM Symposium, Heidelberg, (2007)
14. Ax, A.: The physiological differentiation between fear and anger in humans. Psychosomatic Medicine 55(5), 433--442 (1953)
15. Haag A., Goronzy S., Schaich P., Williams, J.: Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System. In André et al. (eds.): Affective Dialogue Systems, pp 36--48, Springer (2004)
16. Peter, C., Ebert E., Beikirch, H.: A Wearable Multi-Sensor System for Mobile Acquisition of Emotion-Related Physiological Data. In: Proceedings of the 1st International Conference on Affective Computing and Intelligent Interaction, Beijing 2005. Berlin, Heidelberg, New York: Springer (2005)
17. Tan H.Z., Slivovsky L.A., Pentland, A.: A Sensing Chair Using Pressure Distribution Sensors. IEEE/ASME Transactions On Mechatronics, 6(3) (2001)
18. Peter C., Schultz R., Voskamp J., Urban B., Nowack N., Janik H., Kraft K., Göcke, R.: EREC-II in Use – Studies on Usability and Suitability of a Sensor System for Affect Detection and Human Performance Monitoring. In: Jacko, J. (ed.) Human-Computer Interaction, Part III, HCII 2007, LNCS 4552, pp. 465--474. Berlin Heidelberg: Springer (2007)
19. Kanade, T., Cohn, J.F., Tian, Y.: Comprehensive database for facial expression analysis. International Conference on Automatic Face and Gesture Recognition, pp. 46--53 (2000)